

# Magnetic Instability on the Surface of Topological Insulators

Yuval Baum and Ady Stern

*Department of Condensed Matter Physics, Weizmann Institute of Science, Rehovot 76100, Israel*

Gapless surface states that are protected by time reversal symmetry and charge conservation are among the manifestations of 3D topological insulators. In this work we study how electron-electron interaction may lead to spontaneous breaking of time reversal symmetry on surfaces of such insulators. We find that a critical interaction strength exists, above which the surface is unstable to spontaneous formation of magnetization, and study the dependence of this critical interaction strength on temperature and chemical potential.

PACS numbers:

Topological states of matter are one of the main themes of modern condensed matter physics. A recent important discovery in this field are Topological Insulators [1, 2]. In three dimensions Topological Insulators are electronic materials that have a bulk gap like an ordinary insulator, but have conducting states on their surface [1, 3]. The gapless spectrum on the surface arises from the nontrivial topological order of the band structure [4, 5]. It is composed of an odd number of Dirac cones in the Brillouin zone of the two dimensional surface. These Dirac cones cannot be gapped for as long as time reversal symmetry and charge conservation are maintained.

In this work we explore the conditions under which the surface of a strong topological insulator may lower its energy by a spontaneous breaking of time reversal symmetry through the formation of a uniform spin polarization. The way energy may be saved by spin polarization is most easily understood for the case where there is one Dirac cone on each surface, and the chemical potential lies at the Dirac point. For such a case, two sources for energy gain may be identified: first, if the spin polarization introduces a mass term to the Dirac Hamiltonian, the filled electronic states are lowered in energy, while the energy of the unfilled states is raised. And second - as usual for electrons - if all spins are polarized, two electrons cannot reside at the same point, and the short-distance energy cost is reduced. This is not the case in the absence of magnetization. Even though the half filled Dirac cone includes just one electron for every momentum state, the variation of the spin's direction with that of the momentum leads to a non-zero probability of two electrons being at the same point, and thus to a gain of interaction exchange energy by spin polarization.

In our analysis, carried out within the Hartree-Fock approximation, we find that there is a critical interaction strength above which the spin polarized ground state is favorable in energy compared to the non-interacting ground state. We find this to be the case for every value of the chemical potential, for all temperatures much smaller than the bulk energy gap, and for both contact interaction and Coulomb interaction. Of the possible spin polarizations we find that the lowest energy corresponds to an out-of-plane polarization, which introduces an energy gap to the Dirac cone at the Dirac point. For contact interaction, we find the phase transition from the unpolarized to the polarized phase to be of second order both as a function of interaction strength and as a function of temperature. Finally, we find the energy gap to rise sharply as temperature is lowered below the temperature  $T_c$  of the phase transition, and reach values comparable to  $T_c$  at about  $T \approx 0.9T_c$ . We note that a gapped Dirac cone was recently observed by angle resolved photo emission (ARPES) in the topological insulator phase [10], but the information available is not sufficient to judge whether this observation is related to the mechanism we study.

We start with the non-interacting Hamiltonian. ARPES measurements carried out on two topological insulators,  $\text{Bi}_2\text{Te}_3$  and  $\text{Bi}_2\text{Se}_3$ , have found that while near the Dirac point the spectrum is Dirac-like, an hexagonal warping of the Fermi surface occurs away from the Dirac point [6, 7]. This phenomenon was modeled by Fu [8], who suggested an effective Bloch Hamiltonian of the surface including a cubic term:

$$\mathcal{H}(\mathbf{k}) = v_0(k_x\sigma_y - k_y\sigma_x) + \lambda k^3 \cos(3\theta)\sigma_z \quad (1)$$

where  $v_0$  is the electron velocity near the Dirac point,  $\lambda$  is the warping parameter and  $\sigma$  are the Pauli matrices. The above spectrum contains two branches that meet at the Dirac point. This Hamiltonian is time-reversal symmetric. At the time-reversal invariant point  $k = 0$  the spectrum is degenerate and the degeneracy is protected by the time-reversal symmetry.

The interaction part of the Hamiltonian is

$$\mathcal{H}_I = \int d\mathbf{r} d\mathbf{r}' \psi_{\sigma}^{\dagger}(\mathbf{r}) \psi_{\sigma}(\mathbf{r}) V(\mathbf{r}-\mathbf{r}') \psi_{\sigma'}^{\dagger}(\mathbf{r}') \psi_{\sigma'}(\mathbf{r}') \quad (2)$$

Here  $\psi_{\sigma}^{\dagger}(\mathbf{r})$  is the creation operator of an electron at point  $\mathbf{r}$  and spin  $\sigma$ . We will assume that the following operator has a non vanishing expectation value:

$$\langle \psi_{\sigma}^{\dagger}(\mathbf{r}) \psi_{\sigma'}(\mathbf{r}') \rangle \equiv M_{\sigma, \sigma'}(\mathbf{r}-\mathbf{r}') \quad (3)$$

Strictly speaking, we can interpret  $M$  as the magnetization only when  $\mathbf{r} = \mathbf{r}'$ . However, the expectation value (3) breaks time reversal symmetry even when  $\mathbf{r} \neq \mathbf{r}'$ . We confine ourselves to uniform states, in which  $M$  will depend only on  $\mathbf{R} \equiv \mathbf{r} - \mathbf{r}'$ . This case is most relevant when the Fermi-energy is located at the vicinity of the

Dirac point, otherwise spin-wave formation is an interesting possibility.

The expectation value (3) may be written in a matrix form as  $M_{\sigma,\sigma'}(\mathbf{r}-\mathbf{r}') = M_0(\mathbf{R}) \cdot \mathbf{I} + \vec{M}(\mathbf{R}) \cdot \vec{\sigma}$ , but since we are interested only in the spin part of the interaction and in particular the out-of-plane spin ( $z$  direction), we will assume that only  $M_z \neq 0$ , hence:  $M_{\sigma,\sigma'}(\mathbf{r}-\mathbf{r}') \equiv M(\mathbf{R})\sigma_z$ . In-plane magnetization will be discussed in a later part of the paper.

By denoting  $F_{\mathbf{k}} = \int d\mathbf{R} V(\mathbf{R}) M(\mathbf{R}) \exp(-i\mathbf{k} \cdot \mathbf{R})$ , we can write the mean-field Hamiltonian as:

$$H^{MF} \equiv \int d\mathbf{R} V(\mathbf{R}) M^2(\mathbf{R}) + \int \frac{d\mathbf{k}}{(2\pi)^2} c_{\mathbf{k}}^\dagger \hat{h}(\mathbf{k}) c_{\mathbf{k}} \quad (4)$$

where

$$\hat{h} = v_0(k_x\sigma_y - k_y\sigma_x) + (\lambda k^3 \cos 3\theta - 2F_{\mathbf{k}})\sigma_z$$

For a self-consistent determination of the out-of-plane magnetization, we denote the eigenvalues and the corresponding eigenstates of  $\hat{h}$  as  $\epsilon_{\pm}(\mathbf{k})$  and  $(\chi_{\uparrow}, \chi_{\downarrow})_{\pm}$  respectively. The  $\pm$  signs refer to the positive and negative energy states. Hence, the magnetization per unit volume is:

$$m = \frac{1}{\Omega} \sum_{\mathbf{k}} (\langle \sigma_{z-} \rangle f_{\epsilon_-} + \langle \sigma_{z+} \rangle f_{\epsilon_+}) \quad (5)$$

where

$$\langle \sigma_{z\pm} \rangle = |\chi_{\uparrow}|_{\pm}^2 - |\chi_{\downarrow}|_{\pm}^2$$

and  $f_{\epsilon}$  is the Fermi-Dirac function at energy  $\epsilon$ . The trivial solution ( $m = 0$ ) always exists, but we are interested in non trivial solutions of this equation. The total energy per unit volume is:

$$E_m = \int d\mathbf{R} V(\mathbf{R}) M^2(\mathbf{R}) + \frac{1}{\Omega} \sum_{\mathbf{k}} (\epsilon_- f_{\epsilon_-} + \epsilon_+ f_{\epsilon_+}) \quad (6)$$

We will define the energy cost to have finite magnetization as  $\Delta E = E_m - E_0$ . If we find a non trivial solution to Eq. (5) that minimizes  $E_m$  (gives  $\Delta E < 0$ ), we will conclude that the system forms spontaneous surface magnetization.

Although it is hard to determine the exact validity conditions for this variational procedure and to estimate its accuracy, it is still possible to define a necessary condition. Since the procedure above relies on the existence of a bulk gap, it is clear that it requires the interaction energy to be smaller than the bulk gap. We will consider the condition  $U_c < \Delta_{bulk}$  as necessary for the above procedure (where  $U_c$  is the critical interaction energy and  $\Delta_{bulk}$  is the bulk energy gap).

It is simplest to analyze a contact interaction,  $V(\mathbf{R}) = \frac{g}{2}\delta(\mathbf{R})$ . At zero temperature Eq. (5) for the magnetization  $M(0) \equiv m$  becomes:

$$m = \int_{k_F} \frac{d\mathbf{k}}{(2\pi)^2} \frac{gm - \lambda k^3 \cos 3\theta}{\sqrt{v_0^2 k^2 + (gm - \lambda k^3 \cos 3\theta)^2}} \quad (7)$$

where  $d\mathbf{k} = kdkd\theta$  and  $k_F$  is the Fermi-momentum, which is the solution to  $v_0 k_F + \lambda k_F^3 = \mu$ , with  $\mu$  the chemical potential. In general the integration has an upper cutoff in momentum,  $k_{\Lambda}$ , which corresponds to the momentum at which the surface states merge with the three dimensional bands. For  $\lambda = 0$  all results depend linearly on this cutoff, but for  $\lambda \neq 0$  we can identify a momentum scale  $q_0 \equiv (\frac{v_0}{\lambda})^{1/2}$  above which the integrand decays as  $k^{-3}$ . The cutoff dependence then disappears for  $k_{\Lambda} \gg q_0$ .

We find that there is a critical interaction strength  $g_c$  at which a second order phase transition occurs, such that for  $g > g_c$  there is an  $m \neq 0$  solution to (7) with  $\Delta E < 0$ . The critical interaction strength is

$$g_c = \frac{\sqrt{\lambda v_0}}{Y(p_{\Lambda}) - Y(p_F)} = \frac{g_c(\mu = 0)}{1 - \frac{Y(p_F)}{Y(p_{\Lambda})}} \quad (8)$$

where  $p_{\Lambda,F} = k_{\Lambda,F} \sqrt{\frac{\lambda}{v_0}}$  is a dimensionless parameter and the function  $Y$  is defined by:

$$Y(p) = \int_0^p \frac{p' dp'}{(2\pi)^2} \int_0^{2\pi} d\theta \frac{p'^2}{(p'^2 + (p'^3 \cos 3\theta)^2)^{3/2}} \quad (9)$$

The function  $Y(p_{\Lambda})$  is a continuous, positive and monotonically increasing function of  $p_{\Lambda}$  with the following asymptotic behavior:

$$Y(p_{\Lambda}) \sim \begin{cases} \frac{p_{\Lambda}}{2\pi} & , p_{\Lambda} \ll 1 \\ 0.247 & , p_{\Lambda} \gg 1 \end{cases} \quad (10)$$

For  $\mu = 0$  and for  $\lambda \rightarrow 0$  Eq. (8) becomes  $g_c = 2\pi \cdot \frac{v_0}{k_{\Lambda}}$ , while for  $k_{\Lambda} \gg (\frac{v_0}{\lambda})^{1/2}$  Eq. (8) becomes  $g_c \approx 4.1 \cdot \sqrt{v_0 \lambda}$ . The resulting phase diagram ( $\mu = 0$ ) in the large cutoff limit is presented in Fig. (1).

Clearly, for all values of  $\lambda$  and  $k_{\Lambda} > k_F$ , the critical interaction strength  $g_c$  increases as  $|\mu|$  increases. In the limit  $\lambda \rightarrow 0$  Eq. (8) gets a simple form:

$$g_c(\mu) = g_c(\mu = 0) \cdot \frac{1}{1 - \frac{k_F}{k_{\Lambda}}}$$

As expected for  $k_F \rightarrow k_{\Lambda}$  (full band) magnetization formation is not possible. We note, however, that as long as the chemical potential is far from the bulk bands its effect on  $g_c$  is rather weak.

For  $g \sim g_c$  we can estimate the interaction energy per particle by noticing that  $\langle \mathcal{H}_{\mathcal{I}} \rangle_{\text{per-particle}} = \frac{gn_e}{8}$ , where  $n_e$  is the electron density on the surface. We will use ARPES data of  $\text{Bi}_2\text{Te}_3$  and  $\text{Bi}_2\text{Se}_3$  surfaces [7, 9] to compare the critical interaction energy with the bulk energy gap for a typical surface density (table I). In both materials  $U_c < \Delta_{bulk}$ , which puts the mean field results within the necessary condition for the validity of the approximation.

At non-zero temperatures the transition happens when the free energy of the magnetized surface is lower than that of the unmagnetized one.

TABLE I: Bulk gap vs.  $U_c$  for  $n_e = 5 \cdot 10^{13} \text{ cm}^{-2}$ 

	Bi <sub>2</sub> Te <sub>3</sub>	Bi <sub>2</sub> Se <sub>3</sub>
$\Delta_{\text{bulk}}$	0.16 eV	0.35 eV
$U_c$	0.078 eV	0.066 eV

It is useful to note that:  $\langle \sigma_{z-} \rangle = -\langle \sigma_{z+} \rangle \equiv \langle \sigma_z \rangle$ . For  $\mu = 0$ , the equation for  $m$  becomes:

$$m = \int \frac{d\mathbf{k}}{(2\pi)^2} \frac{x(m, \mathbf{k}) \cdot \tanh \frac{\sqrt{v_0^2 k^2 + x_m^2}}{2T}}{\sqrt{v_0^2 k^2 + x_m^2}} \quad (11)$$

where  $T$  is the temperature and  $x(m, \mathbf{k}) = gm - \lambda k^3 \cos 3\theta$ . The free energy difference ( $\Delta f$ ) per unit volume is

$$T \int \frac{d\mathbf{k}}{(2\pi)^2} \log \left( \frac{1 + \cosh \left( \frac{\sqrt{v_0^2 k^2 + x_m^2}}{T} \right)}{1 + \cosh \left( \frac{\sqrt{v_0^2 k^2 + (\lambda k^3 \cos 3\theta)^2}}{T} \right)} \right) \quad (12)$$

It is natural to define the dimensionless temperature scale  $\tau \equiv \frac{T}{E^*}$  where  $E^* = \left( \frac{v_0^3}{\lambda} \right)^{1/2}$  is the characteristic energy scale introduced by hexagonal warping. The resulting phase diagram is presented in Fig. (1).

As expected, the critical interaction strength  $g_c$  increases with increasing temperature. We note however that the effect on  $g_c$  is rather mild. Again, the phase transition is of second order, both for a fixed temperature as a function of  $g$ , and for a fixed  $g$  as a function of temperature.

A similar analysis may be carried out to examine the conditions for in-plane magnetization to be formed. To that end, we assume a magnetization in the  $x-y$  plane, solve the self consistent equation that determines its magnitude and find the conditions under which it leads to an energy gain. As we now show, for typical parameters ( $v_0, \lambda, k_\Lambda$ ) the critical interaction strength to have in-plane magnetization  $g_{c,x-y}$  is smaller than  $g_{c,z}$ , the critical interaction strength to have a magnetization perpendicular to the plane. We carry out the calculation at

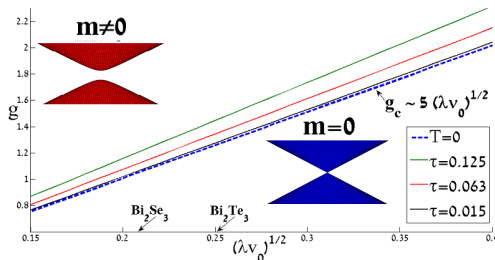


FIG. 1: Phase diagram ( $p_\Lambda = 2$ ) for contact interaction at zero (dashed line) and finite temperatures, where  $\tau \equiv \frac{T}{E^*}$ . A critical interaction exists, beyond which magnetization formation occurs. The critical interaction increases as the temperature increases.

zero temperature. We note that an in-plane magnetization does not create a gap in the spectrum. It does however break time-reversal symmetry and hence removes the topological protection of the Dirac-point.

The Bloch Hamiltonian is:

$$\mathcal{H}(\mathbf{k}) = v_0(k_x \sigma_y - k_y \sigma_x) + \lambda k^3 \cos 3\theta \sigma_z - g \vec{m} \cdot \vec{\sigma} \quad (13)$$

We denote  $(m_x, m_y) \equiv \mathbf{m}_\parallel$ . In the symmetry broken phase, the system will choose a specific direction. Although the system is not rotationally invariant we find that the angle dependence is very small, and we choose  $\mathbf{m}_\parallel = m_x$ . The self consistent equation for  $m_x$  is then given by:

$$m_x = \int_{k_F}^{k_\Lambda} \frac{d\mathbf{k}}{(2\pi)^2} \langle \sigma_x(\mathbf{k}) \rangle \quad (14)$$

where

$$\langle \sigma_x \rangle = \frac{m_x g + v_0 k \sin \theta}{\sqrt{v_0^2 k^2 + (g m_x)^2 + 2 g v_0 m_x k \sin \theta + (\lambda k^3 \cos 3\theta)^2}}$$

All that is left now is to solve Eq. (14), minimize the energy and deduce the critical interaction  $g_{c,x-y}$ . The resulting  $g_{c,x-y}$  has the same form as Eq. (8), with  $Y(p_\Lambda)$  being replaced by  $\Psi(p_\Lambda)$ , where  $\Psi(p_\Lambda)$  is a continuous, positive and monotonically increasing function of  $p_\Lambda$  with the following asymptotic behavior:

$$\Psi(p_\Lambda) \sim \begin{cases} \frac{p_\Lambda}{4\pi} & , p_\Lambda \ll 1 \\ 0.352 & , p_\Lambda \gg 1 \end{cases} \quad (15)$$

The ratio  $\frac{g_{c,x-y}}{g_{c,z}} = \frac{Y(p_\Lambda)}{\Psi(p_\Lambda)}$  is presented in Fig. (2) for  $\mu = 0$ . For  $k_\Lambda < 3.7 \sqrt{\frac{v_0}{\lambda}}$ , the instability to  $m_z$  appears before the instability to in-plane magnetization. Since this value of  $k_\Lambda$  corresponds a warping parameter that is about a hundred times larger than that observed in Bi<sub>2</sub>Te<sub>3</sub>, we conclude that the interesting quantity to examine is the magnetization in the  $z$  direction. This conclusion does not change when the chemical potential is away from the Dirac point.

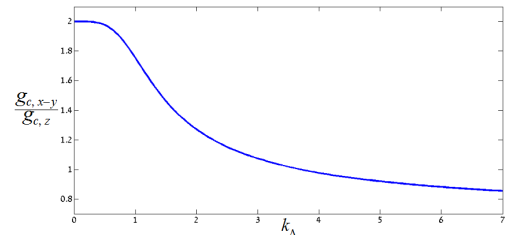


FIG. 2: The ratio between the critical interactions to have in-plane/out-of-plane magnetization as a function of the ratio of  $k_\Lambda$  to the warping wave vector scale  $(\frac{v_0}{\lambda})^{1/2}$ . For typical values the instability to  $m_z$  occurs before the one to  $m_x$ .

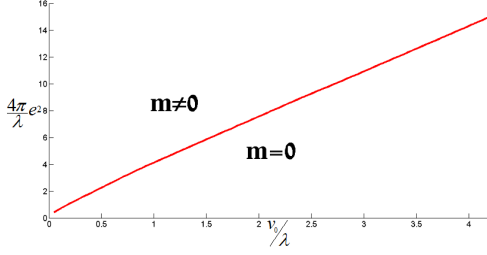


FIG. 3: Phase diagram for Coulomb interaction at zero temperature. A critical interaction exists, beyond which magnetization formation occurs.

Our analysis has so far been focused on contact interaction between the electrons, and has uncovered an instability of the surface to the formation of magnetization, with the instability being strongest for the formation of out-of-plane magnetization at zero temperature and zero chemical potential. The dependence of the critical interaction strength on temperature and chemical potential is however rather weak.

Motivated by these observations we turn to the long range Coulomb interaction,  $V(\mathbf{R}) = \frac{e^2}{R}$ , and study the instability to the formation of out-of-plane magnetization at zero temperature and zero chemical potential [11]. We assume that  $\langle \psi^\dagger_\sigma(\mathbf{r})\psi_{\sigma'}(\mathbf{r}') \rangle \equiv M(\mathbf{R})\sigma_z$  is an exponentially decaying function of  $R$  with a decaying length  $\alpha^{-1}$ . This assumption yields two self consistent equations for the magnetization and for  $\alpha$ :

$$m = \int \frac{d\mathbf{k}}{(2\pi)^2} \frac{\frac{4\pi m e^2}{\sqrt{k^2 + \alpha^2}} - \lambda k^3 \cos 3\theta}{\sqrt{v_0^2 k^2 + \left(\frac{4\pi m e^2}{\sqrt{k^2 + \alpha^2}} - \lambda k^3 \cos 3\theta\right)^2}} \quad (16)$$

$$\frac{m}{\alpha} = \int \frac{d\mathbf{k}}{(2\pi)^2} \frac{\frac{2\pi m e^2}{k\sqrt{k^2 + \alpha^2}} - \lambda k^2 \cos 3\theta}{\sqrt{v_0^2 k^2 + \left(\frac{2\pi m e^2}{k\sqrt{k^2 + \alpha^2}} - \lambda k^2 \cos 3\theta\right)^2}} \quad (17)$$

The trivial solution, in which  $m = 0$  and  $\alpha$  is arbitrary, is always an exact solution to the equations above. After solving numerically for  $m$  and  $\alpha$  and minimizing the energy, we identify the regions of parameter space in which a non-trivial solution exists, as shown in Fig. (3). The values we obtain for  $\alpha^{-1}$  range between 1 and 4 lattice

constants, hence justifying the interpretation of  $m$  as the magnetization.

For the Coulomb interaction the mean interaction energy per particle is of the order of  $e^2 r_0^{-1}$ , where  $r_0$  is the mean distance between electrons on the surface. We can estimate  $r_0 \sim n_e^{-\frac{1}{2}}$ . For a given density we can parameterize the critical interaction strength by defining a critical interaction energy  $U_c = e_c^2 r_0^{-1} \sim e_c^2 n_e^{\frac{1}{2}}$ . Again, for our approximation to be valid we must have  $U_c < \Delta_{bulk}$ . We find this condition to be satisfied for both  $\text{Bi}_2\text{Te}_3$  and  $\text{Bi}_2\text{Se}_3$ : Using Fig. 3 and extracting the density from ARPES data we evaluated the critical interaction strength,  $e_c^2 \sim 0.057e_0^2$  ( $0.072e_0^2$ ) for  $\text{Bi}_2\text{Te}_3$  ( $\text{Bi}_2\text{Se}_3$ ), where  $e_0$  is the electron charge in vacuum. We find the actual Coulomb interaction on the surfaces of these two materials to be too weak for an instability. The dielectric constant on the two surfaces is close to 40, as compared to a value of about 15 for which our approximation leads to an instability.

To summarize, we examined here the possibility that time-reversal symmetry is spontaneously broken on the surface of a three dimensional strong topological insulator due to interactions between surface electrons. We assumed that the surface can be described by a  $2D$  effective Dirac Hamiltonian, and treated interactions within the Hartree-Fock approximation. We found that for a strong enough interaction, both of the contact and the Coulomb types, the surface is unstable to the formation of spontaneous magnetization, with the strongest instability being to the formation of magnetization in the direction perpendicular to the surface. For the contact interaction, we found the transition from the unmagnetized to the magnetized surface to be of second order, both at zero temperature as a function of interaction strength and for a fixed interaction strength as a function of temperature. The dependence of the critical interaction strength on temperature and chemical potential, at least for the contact interaction, is rather mild. We found the interaction strength in the two most studied strong topological insulators to be too weak for an instability, but not by a large factor, making the instability an issue that may be relevant for other topological insulators.

The authors thank the US-Israel Binational Science Foundation, the Minerva foundation and Microsoft's station Q for financial support.

[1] C.L. Kane and M.Z. Hasan, Rev.Mod.Phys.82:3045,2010.  
[2] X.L Qi and S.C Zhang Physics Today, Jan. 2010, Vol. 63, Pg. 33-38  
[3] C.L. Kane and E. J. Mele, Phys. Rev. Lett. 95, 146802, 2005b.  
[4] L. Fu, C. L. Kane, and E. J. Mele, Phys. Rev. Lett. 98, 106803 (2007).  
[5] J.E. Moore and L. Balents, Phys. Rev. B 75, 121306 (2007).

[6] Y. L. Chen *et al.*, Science 325, 178 (2009).  
[7] D. Hsieh *et al.*, Nature 460, 1101 (2009).  
[8] Liang Fu, PRL 103, 266801 (2009).  
[9] K. Kuroda *et al.*, PRL 105, 076802 (2010).  
[10] T. Sato *et al.*, Nature Phys. doi:10.1038/nphys2058 (2011)  
[11] N. M. R. Peres *et al.*, Phys. Rev. B 72, 174406 (2005).